



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) EcoNI (96)

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
PvuII (239) Bsu36I (291)
EcoNI (287)

301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGTTGCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) 501 TCTGTTTTCGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGACTAAAAAATTCTAGCATCTTCCATTTGCCAT
AgeI (552) 1▶ M T K K N S S I F H F A I
BspHI (560)

601 CATCTTTATATTAATACTTGAGATCAGAACTCAATTATCTGATGAAAGTGAATTTTAAATGACAGGTCAAAAAGAGGTCTCACATGTTCCAAAAAC
13▶ I F I L I L E I R T Q L S D E S E F L I D R S K R G L T Y V P K N
701 TTATCCCTGGAAACAACCATCTTAGATATATCATACAACATATTTCTGAGCTTCAGATGCCTGACATCCTCTCACTCAAAGCTGAAGATTTTGATA
47▶ L S L E T T I L D I S Y N Y I S E L Q M P D I L S L S K L K I L I
801 TTTCTATAATAGAAATCCAGTATCTTGACTTGAGTGTTTTTAAATTCACCCAGGAACTGGAATACTTGGATTTGTCCACAACAATTTGGAGAAGATTT
80▶ I S H N R I Q Y L D L S V F K F N Q E L E Y L D L S H N N L E K I S

BspLU11I (983)
901 TTGCCACCCTACTCTGAACCTCAAGCACTTAGACCTCTCATTTAATCCATTTGATGCCCTGCCATATGCCAAGAGTTTGGCAACATGTCTCAACTAGAA
113▶ C H P T L N L K H L D L S F N P F D A L P I C Q E F G N M S Q L E
1001 TTTCTGGGTTGAGTGCCACACAGTTACAGAAATCCAGTGTGACGTCATCCTCATTGACATCAGCAAGTTTTATTGGTCTTAGGAGATACTTATG
147▶ F L G L S A T Q L Q K S S V Q S I T H L H I S K V L L V L G D T Y
1101 GGGAAAGAGAAGATGCCGAGAGCCTTCAAGACCTTAAGACACAGAGTCTGCACATTTTCCACAGGAAAGAAATTCATTTTATTTGGACGTGC
180▶ G E R E D A E S L Q D L K T Q S L H I V F P T G K E F H F I L D V S

XcmI (1208) 1201 AGTCGGCACCACAGTGTGAACTGTCTAATATCAATATGTTGCTTGTGATAATGGGTGTCCTTATTTGAAAAATGTTCTGTCAAACCTCAAAG
BstBI (1270)
Bsp119I (1270)
213▶ V G T T V S L E L S N I K C V L D D N G C P Y F E N V L S K L Q K

ScaI (1394)
1301 AACTCAAGTTATCAAATCTTACTTTAAACAACATTGAAATAACTTGAATTCCTTCTCACGATCCTCCAGTTGGTTTGGCGTACAAACATTGAGTACT
247▶ N S R L S N L T L N N I E I T W N S F F T I L Q L V W R T N I E Y

BstEII (1426) 1401 TCTCCATTTCAAATGTGAAACTACAAGTTACCTTGACTCTAGAGATTTTATTATTCTGACACTTCACTGAAGGCTTGTCTATACAAAAGTTGTCCA
XbaI (1438) 280▶ F S I S N V K L Q G Y L D S R D F D Y S D T S L K A L S I H K V V H
StuI (1472)

1501 TGATGTGTTCACTTCCACAAGTTATGTCTATAAAATATTGCAAATATGAACATCCAGCATCTCACAGTGTCTGCTGCACACATGGTCCACATGGTC
313▶ D V F S L P Q G Y V Y K I L S N M N I Q H L T V S A A H M V H M V

BsrGI (1675)
1601 TGCCCATCCCAAATAGCCATTTCTGTATTTGAATTTTCCAATAATCTCTTAACAGACACAGTTTTATAAAGTGTACAAATTTGGCTAATTTGAAGA
347▶ C P S Q I S P F L Y L N F S N N L L T D T V F I N C T N L A N L K
1701 CACTTATCTACAAAAGAATCAGTTAAAAGAAGTTGTAACATAGTTCAATGACCCAGGAAATGAAGTCTTACAACAAGTGGATGTTAGCCAGAATTC
380▶ T L I L Q K N Q L K E L V N I V H M T Q E M K S L Q Q L D V S Q N S
1801 CCTGATGATGATGAAAGTGAAGAAATGCCCCTTGGCCAGAAATTTAAGTTTAAATATGCTTCAAATATACTTACTGACTCTGTTTTCAGATG
413▶ L M Y D E S E G N C P W A R N L L S L N M S S N I L T D S V F R C

XbaI (1971)
1901 TTACCTCCTCAGATCAAGGTTCTTGATCTTACAATAACAGAATAAGGAGCATCCCTAAAGATGTCAGTGGTCTAGAAACTTTGCAAGAAGTCAACCTTG
447▶ L P P Q I K V L D L H N N R I R S I P K D V T G L E T L Q E L N L

PvuII (2090)
2001 CTTCAATTTCTTAGCCACCTTCTGGATGTGGTATCTTTAGCAGCCTTCCATACTGATCATTGACTATAACTCAATTTCCATCCATCAGCTGATT
480▶ A S N S L A H L P G C G I F S S L S I L I I D Y N S I S N P S A D F

EcoO109I (2122)
2101 CTTCCAGAGCTGCCAGAAGATTAGGTCCCTCAAAGCGGGAACATCCATTCCATGTTCTGTGAGCTAAGAGACTTCACTCAAAGTATAGGCCAAGTA
513▶ F Q S C Q K I R S L K A G N N P F Q C S C E L R D F I Q S I G Q V

BspEI (2247)
2201 TCAAGTGACGTGGTAGAGGGCTGGCCTGAGTCTTATAAGTGTGACTATCCGAAAAGCTACAAGGAAACCCCTCAAAGGACTTCCAGGTATCTGAGCTAT
547▶ S S D V V E G W P E S Y K C D Y P E S Y K G T P L K D F Q V S E L

Acc65I (2396)
2301 CCTGCAACACAGCTCTGCTGATCGTCACCATTGTGGTCCCTGGGCTGGTGTGGTGTGTTGCTGTGACTGTCCTCTGTATCTACCTGGATCGCCCTGGTA
580▶ S C N T A L L I V T I V V P G L V L A V A V T V L C I Y L D L P W Y

Bsu36I (2400)
2401 CCTCAGGATGGTGTGTCAGTGGACCCAGACCCGCGCAGGGCCAGGAATGACCTTGGAAAGAACTCCAAAGAACTCTCCAGTCCATGCTTTTATTTCA
613▶ L R M V C Q W T Q T R R R A R N V P L E E L Q R T L Q F H A F I S
2501 TATAGTGGCAGGATTCTGCCTGGTGAAGAATGAATTAATACCTAAGAAAAGAGATATAAGAAATTTGCTCCATGAGAGAACTTTGTTGCTG
647▶ Y S G H D S A W V K N E L I P N L E K E D I R I C L H E R N F V A
2601 GCAAGAGCATTGTGAAAATATCATCAACTGCATTGAGAAAAGTTACAATCCATCTTTGCTTGTCTCCAACTTTGTCAGAGCGAATGGTGCCATTA
680▶ G K S I V E N I I N C I E K S Y K S I F V L S P N F V Q S E W C H Y

2701 TGAAGCTACTTTGCCACCACAATCTCTCCATGAAGGATCTGATAACTTAATCCTGATCTTCTGGATCCCATCCACAGTATCCATTCCTAGCAGC
713▶ E L Y F A H H N L F H E G S D N L I L I L L D P I P Q Y S I P S S
2801 TACCACAAGCTAAGAGCTCTCATGGCACAGAGAATTATTTGGAATGGCCCAAGGAGAAGAGTAAACACGGACTTTTTGGGCTAACCTAAGAGCATCCA
747▶ Y H K L R A L M A Q R T Y L E W P K E K S K H G L F W A N L R A S

MscI (2947)

2901 TTAATATTAAGCTGATGGAAAAAGCAGCAGAAATACATTAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAT
780▶ I N I K L M E K A A E I H •

NheI (2941)

HpaI (3079) MfeI (3090)

3001 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATT
3101 CATTTTATGTTTCAGGTTTACGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAA
3201 ACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTGCCAATGTGCATTAGCTGTTTGA

SapI (3357)

3301 GCCTCACCTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACAT

SwaI (3428)

EcoO109I (3489)

3401 TCCCTTTTGTAGAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT
3501 AATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTAC
1+1 • N R T Y

3601 TTGAGGGGATGAGTTCCCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGC
135▶ K L P I L E E I T T K V L K G N M E I L V F C D P A Y D S I L E R C
3701 ACATGCCACAGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGATAGGGGTGCCTGACAGCCAAATGGTGTCAAAGTCTTCTGCCGTTGCT
102▶ M G C P S V R I S R D V E D S Y P H R V A V I T D F D K Q G N S

StuI (3853)

3801 CACAGCAGCCCAATGGCAATGGCTTCCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTG
69▶ V A S G I A I A E A C V T V R G I Y A E I H V A S I E G T K T R
3901 ATGGCCGCCCGACATGGTGTCTTGTCTCATAGACATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGG
35▶ I A A G V H H K N D E Y L M T I K E T A V E V L E L D Q Q S I N F T
4001 TCTTCATGGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTA
2▶ K M
4101 TCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTGG

SpeI (4216)

4201 AAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCAGCCCATTTG

SnaBI (4344)

4301 ATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAAT
4401 GCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGCAGTTTACCGTAAATACTC
4501 CACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGC

PacI (4635)

PstI (4628)

SdaI (4627)

BspLU11I (4645)

4601 GGGCCATTACCCTAAGTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTCT
4701 GGGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
4801 TTTCCCTTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTC

ApaLI (4959)

4901 ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCTTATC
5001 CGGTAACATATCGTCTTGTGTCACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGT
5101 GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAG
5201 TTGGTAGCTCTTGTATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA

EagI (5395)

PacI (5375) SwaI (5384) NotI (5394)

5301 TCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCC
5401 GCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAACGAAACAAACAA
5501 ACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA